UDC 551.501.45:551.577.36:519.241.2

# REMARKS ON THOM'S ESTIMATORS FOR THE GAMMA DISTRIBUTION

L. R. SHENTON [1]

Computer Center, The University of Georgia, Athens, Ga.

K. O. BOWMAN [2]

Computing Technology Center, Oak Ridge, Tenn.

### ABSTRACT

Thom's estimators for the two-parameter gamma distribution arise as asymtotic approximations to the maximum likelihood estimators. Being perhaps the simplest estimators known in this case, their properties are here investigated. We show that although they do have slight asymptotic bias even in very large samples, yet for almost the whole of the parameter space they have smaller asymptotic variances than the maximum likelihood estimators; more than this there is evidence that in finite samples the property still holds. As for the type of the sampling distributions involved, Thom's estimators are in general slightly nearer to normality than the maximum likelihood estimators.

The occurrence of estimators that are improvements on the maximum likelihood estimators, be the improvement only slight, is rather rare and becomes of particular interest when they arise in a practical situation.

## 1. INTRODUCTION

It is well known that the gamma distribution with two parameters has wide application in meteorology; for example, it may be used to describe the distribution of rainfall amount per period (day, week, month), also the distribution of rainfall amount per storm (rainy period), and areal rainfall. The smoothed distribution can then be used, among other things, to assess probabilities of precipitation amounts in excess of given values. The relevance of the distribution to these situations has been described by Thom (1957, 1958, 1968) and others.

Again, the gamma distribution is receiving considerable attention in certain aspects of the interpretation of the relation between precipitation and streamflow; references are to be found in connection with frequency analysis in hydrology.

Moreover, because of the feedback from very large sets of data from automated instrumentation in precipitation and water runoff, it is clear that the gamma distribution will receive increased attention, so that a more detailed knowledge of its properties is likely to be useful. When writing the density of the distribution in the form

$$f(x;\gamma,\beta)=\frac{1}{\beta^{\gamma}\Gamma(\gamma)}\,x^{\gamma-1}e^{-x/\beta},\,0<x<\infty,\,(\gamma>0,\beta>0),\quad (1)$$

the maximum likelihood estimators $\hat{\beta}$, $\hat{\gamma}$ of the parameters, based on a random sample $x_1, x_2, \ldots, x_n$, are given by

$$\ln\hat{\gamma}-\psi(\hat{\gamma})=\ln\,(m_1/g)\qquad (2a)$$

and

$$\hat{\gamma}\hat{\beta}=m_1,\qquad (2b)$$

where $\psi(x)=d\,\ln\Gamma(x)/dx$ (the psi-function), and $m_1$ and $g$ are the sample arithmetic and geometric means, respectively. Since equations (2) are not quite easy to solve (there is of course no particular difficulty if one has access to a digital computer), Thom (1957, 1958) has introduced the estimators

$$\overset{*}{\gamma}=\frac{1+\sqrt{(1+4y/3)}}{4y},\qquad (3a)$$

and

$$\overset{*}{\beta}=m_1/\overset{*}{\gamma}\qquad (3b)$$

where $y=\ln(m_1/g)$. He derived these by replacing the psi-function in (2a) by the first three terms of its asymptotic expansion for large $\gamma$. In passing, we note that Greenwood and Durand (1960), a year or so later than Thom, introduced a very accurate rational fraction approximation to $\hat{\gamma}$ in terms of the random variate $y$; it seems fair to say that their approximation is not as simple as Thom's.

In this note, we show that Thom's statistics are:

a) slightly biased, no matter how large the sample; however, this bias is almost negligible for $\gamma>>0$, and indeed is only of any real importance if $\gamma$ is small (say less than 0.1 approximately); the bias in finite samples is about the same as for the maximum likelihood estimators;

b) superior to the maximum likelihood estimators because their variances are less in large sample theory; there is evidence that this property holds in finite samples also;

c) about as near to normality (as measured by skewness and kurtosis) as the maximum likelihood estimators; actually the distribution of $\overset{*}{\beta}$ is generally nearer to the normal form than that of $\hat{\beta}$.

If the remark in b) causes surprise, one may recall that the Cramer-Rao inequality (Cramer 1946), at least in the one-parameter estimation situation, does not preclude the possibility of improving the asymptotic variance, although this may be at the expense of a permanent bias.

Since the development of the results has much in common, from a procedural point of view, with recent

[1] This work was partially supported through the Southeastern Forest Experiment Station with Department of Defense funds, IP R19-65-WEPS.

[2] Part of this paper is based on work performed at the Computing Technology Center, operated by the Union Carbide Corporation, Nuclear Division, for the U.S. Atomic Energy Commission.

work of the authors (for example, Shenton and Bowman 1968, 1969), we shall omit some of the details.

## 2. EXPANSIONS FOR THE ESTIMATORS

It would appear to be obvious that expansions for $\overset{*}{\gamma}$, $\overset{*}{\beta}$ should use $y$ as the variable. However, the derivatives of the estimators with respect to $y$ depend on $n$, the sample size, and this causes rather serious complications when we select the coefficients of powers of $n^{-1}$ in the moments. This difficulty is avoided if we use a bivariate Taylor expansion in the arguments

$$u=m_1-E\,m_1 \quad \text{and} \quad v=\ln g-E\ln g, \tag{4}$$

where $E$ is the expectation operator. Thus with

$$u=m_1-\gamma\beta \text{ and } v=t-\tau \tag{5}$$

where $t=\ln g$, $\tau=Et=\psi(\gamma)+\ln\beta$, we have

$$\overset{*}{\gamma}=\gamma_0+\sum_{s=1}^{\infty}\left(u\frac{\bar\partial}{\partial m_1}+v\frac{\bar\partial}{\partial t}\right)^s\frac{\overset{*}{\gamma}}{s!}, \tag{6a}$$

and

$$\overset{*}{\beta}=\beta_0+\sum_{s=1}^{\infty}\left(u\frac{\bar\partial}{\partial m_1}+v\frac{\bar\partial}{\partial t}\right)^s\frac{\overset{*}{\beta}}{s!}, \tag{6b}$$

where a typical term in (6a), for example, is

$$\binom{s}{r}u^r v^{s-r}\frac{\bar\partial^s\overset{*}{\gamma}}{\partial m_1^r\partial t^{s-r}}, \qquad r=0,1,\ldots,s,$$

in which the derivative factor is to be interpreted as

$$\left.\frac{\partial^s\overset{*}{\gamma}}{\partial m_1^r\partial t^{s-r}}\right|_{m_1=\gamma\beta,\ t=\tau,\ \overset{*}{\gamma}=\gamma_0}$$

The equations (6) will be seen to be consistent in the sense that $u=v=0$ clearly implies $\overset{*}{\gamma}=\gamma_0$, $\overset{*}{\beta}=\beta_0$. Now from

$$y=\ln(u+\gamma\beta)-(v+\tau)$$

for $u=v=0$,

$$y_0=\ln\gamma-\psi(\gamma),$$

from which

$$\gamma_0=\frac{1+\sqrt{(1+4y_0/3)}}{4y_0} \tag{7a}$$

and

$$\beta_0=\frac{\gamma\beta}{\gamma_0}. \tag{7b}$$

These expressions give the *dominant asymptotic terms* in the mean values of $\overset{*}{\gamma}$, $\overset{*}{\beta}$. Observe in passing that $\gamma_0-\gamma$ is related to $\Delta\hat\gamma$ as tabulated by Thom (1958, page 119). How serious are the discrepancies $\gamma_0-\gamma$, $\beta_0-\beta$? Since the estimators were derived by an asymptotic approximation to $\psi(\gamma)$ for large $\gamma$, let us see how they behave in this case. We find, using the asymptotic expansion for the psi-function (Thom 1958, Abramowitz and Stegun 1964),

$$\gamma_0\sim\gamma+\frac{1}{60\gamma^2}-\frac{1}{180\gamma^3}-\frac{23}{3780\gamma^4}+\frac{523}{226800\gamma^5}+\cdots \tag{8a}$$

and

$$\beta_0\sim\beta\left(1-\frac{1}{60\gamma^3}+\frac{1}{180\gamma^4}+\frac{23}{3780\gamma^5}-\frac{23}{11340\gamma^6}+\cdots\right) \tag{8b}$$

where the symbol $\sim$ refers to asymptotic equivalence (de Bruijn 1961). Evidently, the permanent bias is quite small as long as $\gamma$ is not small; in this case we require the expansion (Abramowitz and Stegun 1964, using 6.3.12 in conjunction with 6.3.5)

$$\psi(\gamma)=-\frac{1}{\gamma}-C-\sum_{m=2}^{\infty}(-1)^m\zeta(m)\gamma^{m-1},\ (|\gamma|<1), \tag{9}$$

to obtain

$$y_0=\frac{1}{\gamma}+\ln\gamma-C+\sum_{m=2}^{\infty}(-1)^m\zeta(m)\gamma^{m-1} \tag{10}$$

in which $C=0.5772\ldots$ is Euler's constant and $\zeta(m)=\sum_{s=1}^{\infty}m^{-s}$ is the Rieman zeta function. Reverting to (7) we have

$$\gamma_0\sim\sqrt{(\gamma/6)} \quad\text{and}\quad \beta_0\sim\beta\sqrt{(6\gamma)},\ (\gamma\text{ small and positive}). \tag{11}$$

Thus in this case, the estimators do not even have the correct order of magnitude in large samples; however, this is not particularly important since the magnitudes involved are small, and in any case one must recall from the method the estimators were conceived that it would indeed be surprising if they had good properties for small values of $\gamma$. We shall give numerical assessments of the biases in infinite and finite samples in the sequel.

## 3. THE ASYMPTOTIC BIASES AND COVARIANCES

We show briefly how to set up the $n^{-1}$ coefficients in the biases and covariances. For these we need the joint moments of $u$, $v$, and the bivariate derivatives of $\overset{*}{\gamma}$, $\overset{*}{\beta}$ with respect to $m_1$, $t$, taken at the population point.

The moments can be found from the joint moment generating function (Shenton and Bowman 1968)

$$E\,e^{pu+qv}=E\,e^{p(m_1-\gamma\beta)+q(\ln g-\tau)}=e^{-p\gamma\beta-q\tau}(E\,e^{px/n+q\ln x/n})^n \tag{12}$$

because of statistical independence. But from equation (1),

$$E\,x^{q/n}e^{px/n}=\frac{\beta^{q/n}\Gamma(\gamma+q/n)}{(1-p\beta/n)^{\gamma+q/n}\Gamma(\gamma)},$$

so that

$$E\,e^{pu+qv}=\frac{e^{-p\gamma\beta-q\tau}\beta^q}{(1-p\beta/n)^{n\gamma+q}}\left\{\frac{\Gamma(\gamma+q/n)}{\Gamma(\gamma)}\right\}^n. \tag{13}$$

From this, the joint cumulant generating function may be expanded as

$$\phi(p,q)=(q+n\gamma)\sum_{s=2}^{\infty}\frac{1}{s}\left(\frac{p\beta}{n}\right)^s+\frac{pq\beta}{n}+n\sum_{s=2}^{\infty}\frac{1}{s!}\left(\frac{q}{n}\right)^s\psi_{s-1}(\gamma) \tag{14}$$

where

$$\psi_s(\gamma)=d^s\psi(\gamma)/d\gamma^s,\quad s=1,2,\ldots .$$

Hence, for the joint cumulants of $u$, $v$, we have

$$\begin{aligned}\kappa_{r,s}&=\frac{\gamma(r-1)!\beta^r}{n^{r-1}}, && s=0, r\geq 2;\\ &=\frac{(r-1)!\beta^r}{n^r}, && s=1, r\geq 1;\\ &=\frac{\psi_{s-1}(\gamma)}{n^{s-1}}, && r=0, s\geq 1;\\ &=0, \text{ otherwise.}\end{aligned}\tag{15}$$

The joint central moments of $u$, $v$, defined by

$$\mu_{r,s}=Eu^rv^s$$

can now be evaluated recursively (and to a high order on a digital computer) from expression (8) in Shenton and Bowman (1968). As examples we have

$$\mu_{2,0}=\beta^2\gamma/n,$$
$$\mu_{1,1}=\beta/n,$$

and

$$\mu_{0,2}=\psi_1(\gamma)/n.$$

Turning to the derivatives, we have from (3a)

$$\frac{1}{2\overset{*}{\gamma}}+\frac{1}{12\overset{*}{\gamma}{}^2}=\ln m_1-t,\tag{16}$$

so that writing $\overset{*}{\gamma}_{1,0}$, $\overset{*}{\gamma}_{0,1}$, (with obvious extensions) for the partial derivatives with respect to $m_1$, $t$ respectively

$$\left\{\begin{aligned}&\left(\frac{1}{2\overset{*}{\gamma}{}^2}+\frac{1}{6\overset{*}{\gamma}{}^3}\right)\overset{*}{\gamma}_{1,0}=\frac{1}{m_1}\\ &\left(\frac{1}{2\overset{*}{\gamma}{}^2}+\frac{1}{6\overset{*}{\gamma}{}^3}\right)\overset{*}{\gamma}_{0,1}=1.\end{aligned}\right.\tag{17}$$

Note the simple relation

$$m_1\overset{*}{\gamma}_{1,0}+\overset{*}{\gamma}_{0,1}=0$$

which is the analogue of (16) in Shenton and Bowman (1968) and which is of considerable use in setting up higher derivatives.

Similarly for the second-order derivatives,

$$\begin{aligned}&H_2\overset{*}{\gamma}_{2,0}-H_3\overset{*}{\gamma}{}^2_{1,0}=\frac{1}{m_1^2},\\ &H_2\overset{*}{\gamma}_{1,1}-H_3\overset{*}{\gamma}_{1,0}\overset{*}{\gamma}_{0,1}=0,\\ \text{and}\quad &H_2\overset{*}{\gamma}_{0,2}-H_3\overset{*}{\gamma}{}^2_{0,1}=0\end{aligned}\tag{18}$$

where

$$H_2=\frac{1}{2\overset{*}{\gamma}{}^2}+\frac{1}{6\overset{*}{\gamma}{}^3},$$

and

$$H_3=\frac{1}{\overset{*}{\gamma}{}^3}+\frac{1}{2\overset{*}{\gamma}{}^4}.$$

The population values of the derivatives, $\gamma_{1,0}$, $\gamma_{0,1}$ and so on, follow from (18), and so for the mean value of $\overset{*}{\gamma}$, as far as the $n^{-1}$ coefficient goes (say $E_1\overset{*}{\gamma}$), we have

$$E_1\overset{*}{\gamma}=\frac{1}{2!}\left(\mu_{2,0}\gamma_{2,0}+2\mu_{1,1}\gamma_{1,1}+\mu_{0,2}\gamma_{0,2}\right)$$

$$=\frac{1}{2h_2}\left\{\gamma\beta^2\left(h_3\gamma_{1,0}^2+\frac{1}{\gamma^2\beta^2}\right)+2\beta h_3\gamma_{1,0}\gamma_{0,1}+h_3\gamma_{0,1}^2\psi_1(\gamma)\right\},$$

where $h_2$, $h_3$ are the population values of $H_2$, $H_3$, respectively. Thus after simplification and the use of (17), we find

$$E_1\overset{*}{\gamma}=\frac{1}{2}\left(\psi_1(\gamma)-\frac{1}{\gamma}\right)\frac{\left(\frac{1}{\gamma_0^3}+\frac{1}{2\gamma_0^4}\right)}{\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)^3}+\frac{1}{2\gamma\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)}.\tag{19a}$$

Similarly by differentiating the relation $\overset{*}{\gamma}\,\overset{*}{\beta}=m_1$ with respect to $m_1$ and $t$ to set up the second-order derivatives of $\overset{*}{\beta}$ in terms of those for $\overset{*}{\gamma}$, we find after simplification

$$E_1\overset{*}{\beta}=-\frac{\beta_0}{\gamma_0}E_1\overset{*}{\gamma}+\frac{\beta_0}{\gamma_0}\frac{\left(\psi_1(\gamma)-\frac{1}{\gamma}\right)}{\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)^2}.\tag{19b}$$

It is interesting to note that since asymptotically

$$\frac{1}{2\gamma^2}+\frac{1}{6\gamma^3}\sim\psi_1(\gamma)-\frac{1}{\gamma},\quad(\gamma\gg 0)$$

and

$$\frac{1}{\gamma^3}+\frac{1}{2\gamma^4}\sim-\psi_2(\gamma)-\frac{1}{\gamma^2},$$

the $n^{-1}$ biases of Thom's estimators are from (19) approximately

$$E_1\overset{*}{\gamma}\sim\tfrac{1}{2}(-2+\gamma\psi_1(\gamma)-\gamma^2\psi_2(\gamma))/(\gamma\psi_1(\gamma)-1)^2$$

and

$$E_1\overset{*}{\beta}\sim\frac{\beta}{2}\left(\psi_1(\gamma)+\gamma\psi_2(\gamma)\right)/(\gamma\psi_1(\gamma)-1)^2$$

which are exactly those for the maximum likelihood estimators (see (18.1) and (18.2) of Shenton and Bowman 1968).

The $n^{-1}$ covariances, denoted by $\text{var}_1\overset{*}{\gamma}$ etc., are derived similarly, and we find

$$\text{var}_1\,\overset{*}{\gamma}=\frac{\psi_1(\gamma)-\frac{1}{\gamma}}{\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)^2},\tag{20a}$$

$$\text{cov}_1\,(\overset{*}{\gamma},\overset{*}{\beta})=-\beta_0\frac{\left(\psi_1(\gamma)-\frac{1}{\gamma}\right)}{\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)^2},\tag{20b}$$

and

$$\text{var}_1\,\overset{*}{\beta}=\frac{\beta^2\gamma}{\gamma_0^2}+\left(\frac{\beta_0}{\gamma_0}\right)^2\frac{\left(\psi_1(\gamma)-\frac{1}{\gamma}\right)}{\left(\frac{1}{2\gamma_0^2}+\frac{1}{6\gamma_0^3}\right)^2}.\tag{20c}$$

TABLE 1.—*Biases for the maximum likelihood and Thom's estimators ($n^{-1}$ through $n^{-6}$ terms)*

| $\gamma$ | | $n^{-1}$ | $n^{-2}$ | $n^{-3}$ | $n^{-4}$ | $n^{-5}$ | $n^{-6}$ |
|---|---|---|---|---|---|---|---|
| | | (a) $E(\hat{\gamma}-\gamma)/\gamma$ | (b) $E(\overset{*}{\gamma}-\gamma)/\gamma$ | | | | |
| 0.1 | (a) | 1.684 | 4.114 | 9.231 | 14.35 | 61.70 | 577.6 |
| | (b) | 1.612 | 3.885 | 8.797 | 14.08 | 62.93 | 580.9 |
| 0.5 | (a) | 2.171 | 6.001 | 17.82 | 54.42 | 164.3 | 479.3 |
| | (b) | 2.122 | 6.005 | 17.88 | 54.29 | 163.8 | 482.0 |
| 1.0 | (a) | 2.463 | 7.236 | 21.81 | 65.40 | 195.0 | 587.2 |
| | (b) | 2.450 | 7.246 | 21.79 | 65.39 | 195.2 | 586.3 |
| 5.0 | (a) | 2.871 | 8.610 | 25.83 | 77.48 | 232.5 | 697.3 |
| | (b) | 2.871 | 8.610 | 25.83 | 77.48 | 232.5 | 697.2 |
| | | (a) $E(\hat{\beta}-\beta)/\beta$ | (b) $E(\overset{*}{\beta}-\beta)/\beta$ | | | | |
| 0.1 | (a) | −0.591 | −0.951 | −0.057 | 9.064 | −0.389 | −432.5 |
| | (b) | − .441 | − .558 | − .077 | 4.771 | −1.841 | −225.9 |
| 0.5 | (a) | −0.808 | −0.278 | 0.069 | 0.184 | −0.390 | −0.959 |
| | (b) | − .831 | − .202 | .081 | .049 | − .198 | .206 |
| 1.0 | (a) | −0.913 | −0.114 | 0.065 | −0.120 | 0.114 | 0.823 |
| | (b) | − .934 | − .071 | .027 | − .040 | .053 | .027 |
| 5.0 | (a) | −0.995 | −0.001 | −0.005 | 0.002 | −0.001 | 0.000 |
| | (b) | − .996 | − .001 | − .003 | .001 | − .000 | .153 |

TABLE 2.—*Variances for maximum likelihood and Thom's estimators ($n^{-1}$ through $n^{-6}$ terms)*

| $\gamma$ | | $n^{-1}$ | $n^{-2}$ | $n^{-3}$ | $n^{-4}$ | $n^{-5}$ | $n^{-6}$ |
|---|---|---|---|---|---|---|---|
| | | (a) var $(\hat{\gamma}/\gamma)$ | (b) var $(\overset{*}{\gamma}/\gamma)$ | | | | |
| 0.1 | (a) | 1.094 | 7.240 | 37.47 | 175.9 | 806.8 | 3697 |
| | (b) | 1.078 | 6.513 | 33.34 | 159.4 | 753.9 | 3545 |
| 0.5 | (a) | 1.363 | 12.43 | 84.93 | 517.2 | 2952 | 16079 |
| | (b) | 1.272 | 12.03 | 83.94 | 515.3 | 2944 | 16040 |
| 1.0 | (a) | 1.551 | 15.77 | 114.4 | 714.9 | 4105 | 22416 |
| | (b) | 1.513 | 15.67 | 114.2 | 714.1 | 4103 | 22412 |
| 5.0 | (a) | 1.876 | 20.47 | 152.1 | 959.2 | 5540 | 30380 |
| | (b) | 1.875 | 20.47 | 152.1 | 959.2 | 5540 | 30383 |
| | | (a) var $(\hat{\beta}/\beta)$ | (b) var $(\overset{*}{\beta}/\beta)$ | | | | |
| 0.1 | (a) | 11.09 | −1.387 | −22.49 | −8.473 | 227.8 | 263.4 |
| | (b) | 5.664 | −3.399 | −6.986 | 0.236 | 79.29 | 4.755 |
| 0.5 | (a) | 3.363 | −1.287 | −2.230 | −0.808 | 1.831 | 3.618 |
| | (b) | 2.940 | −1.713 | −1.521 | − .233 | 1.595 | −0.904 |
| 1.0 | (a) | 2.551 | −1.486 | −1.263 | 0.355 | 0.188 | −4.552 |
| | (b) | 2.440 | −1.703 | − .927 | .301 | −1.151 | −0.484 |
| 5.0 | (a) | 2.076 | −1.955 | −0.093 | −0.048 | 0.041 | −0.053 |
| | (b) | 2.074 | −1.966 | − .096 | − .017 | .007 | .006 |

As was the case for the biases, if we use the approximations for the denominators for large $\gamma$, then the above covariances will be seen to be asymptotically equivalent to the maximum likelihood covariances (Masuyama and Kuroiwa 1951, Shenton and Bowman 1968, Thom 1958). That the variances (20a), (20c) are in general ($\gamma \geq 0.1$ approximately) less than those for the maximum likelihood estimators seems difficult to prove analytically. However after some algebra, it turns out that for large $\gamma$, we have

$$\sqrt{\frac{\mathrm{var}_1\, \overset{*}{\gamma}}{\mathrm{var}_1\, \hat{\gamma}}} \sim 1 - \frac{1}{30\gamma^3} + \frac{1}{60\gamma^4} + \frac{23}{945\gamma^5} + \cdots \tag{21a}$$

and

$$\sqrt{\frac{\mathrm{var}_1\, \overset{*}{\beta}}{\mathrm{var}_1\, \hat{\beta}}} \sim 1 - \frac{1}{15\gamma^3} + \frac{19}{360\gamma^4} + \frac{107}{5040\gamma^5} + \cdots. \tag{21b}$$

These expressions show that for sufficiently large $\gamma$ Thom's estimators are improvements on the maximum likelihood estimators. A few numerical comparisons are given in the column headed $n^{-1}$ in table 2. There is little between the variances if $\gamma \geq 1$, and it is indeed remarkable that the improvement is somewhat more emphatic for $0.1 \leq \gamma < 1.0$. Again notice that the improvement is more marked for $\overset{*}{\beta}$ than for $\overset{*}{\gamma}$.

Since the moments of the estimators considered here involve infinite series in powers of $n^{-1}$, it is conceivable that any advantage enjoyed by one estimator over another for the dominant terms might be lost when higher order terms are included, so that in finite samples the advantage might not hold. We now consider this aspect of the problem.

## 4. BIASES AND COVARIANCES IN FINITE SAMPLES

Using the methods described in Shenton and Bowman (1968) for the maximum likelihood estimators, we have programmed the procedure for evaluating the first four moments of $\overset{*}{\gamma}$ and $\overset{*}{\beta}$, including terms up to order six in $n^{-1}$. A considerable economy in the computations arises because of the scale-free nature of the estimators; thus it is not difficult to show that the joint moments of $\overset{*}{\gamma}$ and $\overset{*}{\beta}/\beta$ are independent of $\beta$. Actually it is beneficial to use $\overset{*}{\gamma}/\gamma$ instead of $\overset{*}{\gamma}$, the division by $\gamma$ tending to stabilize the moments. A selection of the results is given in tables 1 and 2, these referring to the biases and variances; the corresponding results for the maximum likelihood estimators are also included for comparison. It will be noticed that the coefficients of powers of $n^{-1}$ for the bias and variance of $\overset{*}{\gamma}$ are all positive and in general form an increasing sequence for given $\gamma$; by contrast, those for $\overset{*}{\beta}$ oscillate and become small in magnitude for the higher orders. For the most part, the coefficients for the moments of $\overset{*}{\gamma}$ are less than those of $\hat{\gamma}$; this breaks down for the higher order terms in the bias when $\gamma$ is small, although the difference is small. It is not so easy to compare the terms in the moments of $\overset{*}{\beta}$ and $\hat{\beta}$; but in general they are very close in value, and it seems fair to say that Thom's statistic, as far as the $n^{-1}$ and $n^{-2}$ terms are concerned, has the advantage.

Good approximations to the moments of the estimators can be found by noticing the structure of the series involved. Thus for $\overset{*}{\beta}$, since the terms damp off quite rapidly, quite small values of $n$ can be used; however, the minimum sample size required tends to increase with the order of the moment. It turns out that a sample of size

TABLE 3.—*Comparison of standard deviations for maximum likelihood and Thom's estimators in finite samples*

| $\gamma$ | Sample size ($n$) | ML $\sigma(\hat{\gamma})/\gamma$ | Thom $\sigma(\overset{*}{\gamma})/\gamma$ | ML $\sigma(\hat{\beta})/\beta$ | Thom $\sigma(\overset{*}{\beta})/\beta$ |
|---|---|---|---|---|---|
| 0.1 | 25 | 0.2414 | 0.2370 | 0.6634 | 0.4698 |
| | 50 | .1584 | .1564 | .4703 | .3345 |
| | 100 | .1082 | .1071 | .3328 | .2373 |
| 0.5 | 25 | .2856 | .2779 | .3638 | .3388 |
| | 50 | .1817 | .1761 | .2583 | .2410 |
| | 100 | .1223 | .1184 | .1830 | .1710 |
| 1.0 | 25 | .3114 | .3086 | .3155 | .3079 |
| | 50 | .1959 | .1938 | .2245 | .2193 |
| | 100 | .1312 | .1297 | .1592 | .1556 |
| 5.0 | 25 | .3474 | .3474 | .2827 | .2825 |
| | 50 | .2170 | .2170 | .2018 | .2017 |
| | 100 | .1448 | .1448 | .1434 | .1433 |

TABLE 4.—*Skewness ($\sqrt{\beta_1}$) and kurtosis ($\beta_2$) for the maximum likelihood and Thom's estimators*

| $\gamma$ | Sample size ($n$) | ML $\sqrt{\beta_1(\gamma)}$ | Thom $\sqrt{\beta_1(\gamma)}$ | ML $\beta_2(\gamma)$ | Thom $\beta_2(\gamma)$ |
|---|---|---|---|---|---|
| 0.1 | 25 | 1.009 | 0.938 | 5.110 | 4.873 |
| | 50 | 0.658 | .608 | 3.847 | 3.743 |
| | 100 | .449 | .414 | 3.389 | 3.339 |
| 5.0 | 25 | 1.269 | 1.304 | 6.398 | 6.585 |
| | 50 | 0.796 | 0.817 | 4.257 | 4.324 |
| | 100 | .535 | .549 | 3.556 | 3.584 |
| 1.0 | 25 | 1.355 | 1.374 | 6.842 | 6.932 |
| | 50 | 0.848 | 0.861 | 4.418 | 4.455 |
| | 100 | .569 | .578 | 3.625 | 3.642 |
| 5.0 | 25 | 1.403 | 1.404 | 7.075 | 7.077 |
| | 50 | 0.881 | 0.881 | 4.514 | 4.515 |
| | 100 | .592 | .592 | 3.669 | 3.669 |
| | | $\sqrt{\beta_1(\beta)}$ | | $\beta_2(\beta)$ | |
| 0.1 | 25 | 1.526 | 1.408 | 6.808 | 6.136 |
| | 50 | 1.052 | 0.980 | 4.791 | 4.512 |
| | 100 | 0.734 | .688 | 3.868 | 3.742 |
| 0.5 | 25 | 0.919 | 0.864 | 4.418 | 4.229 |
| | 50 | .639 | .602 | 3.682 | 3.595 |
| | 100 | .448 | .423 | 3.334 | 3.293 |
| 1.0 | 25 | 0.772 | 0.743 | 3.992 | 3.906 |
| | 50 | .538 | .519 | 3.480 | 3.440 |
| | 100 | .377 | .364 | 3.236 | 3.127 |
| 5.0 | 25 | 0.612 | 0.611 | 3.582 | 3.579 |
| | 50 | .428 | .427 | 3.284 | 3.283 |
| | 100 | .301 | .300 | 3.140 | 3.140 |

$n=10$ is quite adequate to control the terms in the first four moments of $\overset{*}{\beta}$ for $\gamma \geq 0.1$.

The terms in the series for the moments of $\overset{*}{\gamma}$ being positive and nondecreasing require a different approach; in fact they are so close to the maximum likelihood moments that we used extrapolatory methods on them as is described in Shenton and Bowman (1968).

A comparison of the variances of the two sets of statistics in finite samples is given in table 3. It will be seen that for the parameter values shown, Thom's estimators have smaller variance than the maximum likelihood estimators. The improvement is only slight for the $\gamma$ estimator, but it is more marked for the $\beta$ estimator, especially for small values of $\gamma$. Again, the relationship of the variances for finite samples is about the same as that for the asymptotic variances, that is, Thom's estimators still have smaller variances in general ($\gamma \geq 0.1$).

## 5. SKEWNESS AND KURTOSIS

Following the method soutlined in section 4, the skewness ($\sqrt{\beta_1}=\mu_3/\mu_2^{3/2}$) and kurtosis ($\beta_2=\mu_4/\mu_2^2$) for the distributions of $\overset{*}{\gamma}$ and $\overset{*}{\beta}$ have been computed, and a few values are given in table 4 along with the corresponding values for the maximum likelihood estimators. For the $\beta$-variate, Thom's estimator (for the parameter values considered) is less skew and has a smaller kurtosis; that is, it is nearer the normal distribution. For the $\gamma$-variate, Thom's estimator has a nearer to normal distribution than the maximum likelihood for small values of $\gamma$, but is not quite so good for $\gamma \geq 0.5$ (approximately), although the difference is in general quite small. For large $\gamma$, there is very little to choose between the estimators.

## 6. INDEPENDENT STATISTICS AND THE ESTIMATORS

We observe that $\overset{*}{\beta}=m_1/\overset{*}{\gamma}$, where $\overset{*}{\gamma}$ is a function of $y=\ln(m_1/g)$. It turns out that $m_1$ and $y$ are statistically independent. A proof of this starts with the joint generating function

$$E\, e^{-(p+r)m_1+q\ln g}=\frac{\beta^q}{\left(1+(p+r)\dfrac{\beta}{n}\right)^{n\gamma+q}}\left\{\frac{\Gamma\left(\gamma+\dfrac{q}{n}\right)}{\Gamma(\gamma)}\right\}^n \tag{22}$$

which is assumed to exist. Hence assuming that the operations of integration and expectation can be interchanged, we have

$$\int_0^\infty r^{q-1}\{E\, e^{-(p+r)m_1+q\ln g}\}dr=\Gamma(q)E\, e^{-pm_1-q\ln(m_1/g)}, \tag{23a}$$

and

$$\int_0^\infty \frac{r^{q-1}dr}{\left(1+(p+r)\dfrac{\beta}{n}\right)^{n\gamma+q}}=\left(\frac{n}{\beta}\right)^q \frac{1}{\left(1+p\,\dfrac{\beta}{n}\right)^{n\gamma}}\int_0^1 y^{q-1}(1-y)^{n\gamma-1}dy$$

$$=\left(\frac{n}{\beta}\right)^q \frac{1}{\left(1+p\,\dfrac{\beta}{n}\right)^{n\gamma}}\,\frac{\Gamma(q)\Gamma(n\gamma)}{\Gamma(q+n\gamma)}. \tag{23b}$$

Hence,

$$E\, e^{-pm_1-qy}=\frac{n^q\Gamma(n\gamma)}{\Gamma(n\gamma+q)}\left\{\frac{\Gamma\left(\gamma+\dfrac{q}{n}\right)}{\Gamma(\gamma)}\right\}^n\frac{1}{\left(1+p\,\dfrac{\beta}{n}\right)^{n\gamma}} \tag{24}$$

which is readily seen to factor into $Ee^{-pm_1}$ and $Ee^{-qy}$. Hence, $m_1$ and $y$ are statistically independent.

We can apply this result to simplify the moments of $\overset{*}{\beta}$, for

$$\overset{*}{\beta}=\frac{4m_1y}{1+\sqrt{(1+4y/3)}}=3m_1\{\sqrt{(1+4y/3)}-1\}. \tag{25}$$

Hence let

$$\Phi = E\sqrt{(1+4y/3)}.$$

Then

$$E\overset{*}{\beta} = 3Em_1E\{\sqrt{(1+4y/3)}-1\} = 3\beta\gamma(\Phi-1), \quad (26a)$$

for if two statistics are independent then so are well-behaved functions of them. Again,

$$E\overset{*}{\beta}{}^2 = 9Em_1^2E(2+4y/3-2\sqrt{(1+4y/3)})$$
$$= 9(\gamma^2\beta^2+\gamma\beta^2/n)(2+4y_0/3-2\Phi) \quad (26b)$$

where from (22)

$$y_0 = Ey = \psi(n\gamma)-\psi(\gamma)-\ln n.$$

Hence eliminating $\Phi$, we have the relation

$$E\frac{\overset{*}{\beta}{}^2/9}{(\beta^2\gamma^2+\beta^2\gamma/n)}+\frac{2E\overset{*}{\beta}}{3\beta\gamma}=\frac{4}{3}\{\psi(n\gamma)-\psi(\gamma)-\ln n\}, \quad (27a)$$

which may be written in the form

$$\operatorname{var}\overset{*}{\beta}+(E\overset{*}{\beta})^2+6(\beta\gamma+\beta/n)E\overset{*}{\beta}$$
$$=12(\beta^2\gamma^2+\beta^2\gamma/n)\{\psi(n\gamma)-\psi(\gamma)-\ln n\}. \quad (27b)$$

Further relations involving higher moments could be set up similarly, and they may be used as a check on numerical computations of the moments in finite samples. For example, with $n=25$, $\gamma=0.1$, and $\beta=1$, our computations lead to the value 10.5449 for the left-hand side of (27a), whereas the right-hand side equals 10.5440.

## 7. THE HIGHER ASYMPTOTIC MOMENTS OF $\overset{*}{\gamma}$

Since we know the cumulant generating function of $y$ from (12) or (24), and since the range of $y$ is nonnegative, we can invert the generating function by using the Laplace transform inversion theorem. In this way, the density of $y$ can be written

$$\chi(y)=\frac{1}{2\pi i}\int_{\epsilon-i\infty}^{\epsilon+i\infty} e^{\alpha y+\theta(\alpha)}d\alpha, \quad \epsilon>0, \quad (28)$$

where

$$\theta(\alpha)=\Theta\left(\frac{\partial}{\partial\gamma}\right)y_0',$$

with

$$\Theta(\delta)=\alpha+\frac{\alpha^2}{2!\,n}\delta+\frac{\alpha^3}{3!\,n^2}\delta^2+\ldots,$$

and

$$y_0'=\psi(\gamma)-\psi(n\gamma)+\ln n.$$

Now using the asymptotic expansion for $y_0'$, namely

$$y_0'\sim-\frac{1}{2\gamma}\left(1-\frac{1}{n}\right)-\frac{B_2}{2\gamma^2}\left(1-\frac{1}{n^2}\right)-\frac{B_4}{4\gamma^4}\left(1-\frac{1}{n^4}\right)+\ldots$$

where $B_2=1/6$, $B_4=-1/30$, $B_6=1/42$, etc. are Bernoulli numbers, we have

$$\theta(\alpha)=n\left\{-\frac{1}{2}\left(1-\frac{1}{n}\right)+\frac{B_2}{2\cdot1!}\left(1-\frac{1}{n^2}\right)\frac{\partial}{\partial\gamma}+\frac{B_4}{4\cdot3!}\left(1-\frac{1}{n^4}\right)\frac{\partial^3}{\partial\gamma^3}+\frac{B_6}{6\cdot5!}\left(1-\frac{1}{n^6}\right)\frac{\partial^5}{\partial\gamma^5}+\ldots\right\}\ln\left(1+\frac{\alpha}{n\gamma}\right). \quad (29)$$

But

$$\left(\frac{\partial}{\partial\gamma}\right)^s\ln\left(1+\frac{\alpha}{n\gamma}\right)=-\frac{1}{\gamma}+\frac{1}{\gamma+\alpha/n}, \quad s=1,$$
$$=\frac{1}{\gamma^2}-\frac{1}{(\gamma+\alpha/n)^2}, \quad s=2,$$
$$=\frac{-2!}{\gamma^3}+\frac{2!}{(\gamma+\alpha/n)^3}, \quad s=3,$$

and so on. We thus find

$$\theta(\alpha)=-\frac{1}{2}(n-1)\ln\left(1+\frac{\alpha}{n\gamma}\right)+\omega(\alpha), \quad (30)$$

where

$$\omega(\alpha)=n\sum_{s=1}^{\infty}\frac{B_{2s}}{2s(2s-1)}\left(1-\frac{1}{n^{2s}}\right)\left(-\frac{1}{\gamma^{2s-1}}+\frac{1}{(\gamma+\alpha/n)^{2s-1}}\right),$$

and so

$$e^{\omega(\alpha)}=1+\frac{B_2}{2\gamma}\left(n-\frac{1}{n}\right)\left(-1+\frac{1}{Z}\right)+\frac{B_4}{4\cdot3\gamma^3}\left(n-\frac{1}{n^3}\right)\left(-1+\frac{1}{Z^3}\right)$$
$$+\frac{1}{2!}\left\{\frac{B_2}{2\gamma}\left(n-\frac{1}{n}\right)\left(-1+\frac{1}{Z}\right)+\frac{B_4}{4\cdot3\gamma^3}\left(n-\frac{1}{n^3}\right)\right.$$
$$\left.\times\left(-1+\frac{1}{Z^3}\right)+\ldots\right\}^2+\frac{1}{3!}\left\{\frac{B_2}{2\gamma}\left(n-\frac{1}{n}\right)\right.$$
$$\left.\times\left(-1+\frac{1}{Z}\right)+\ldots\right\}^3+\ldots \quad (31)$$

where $Z=1+\alpha/n\gamma$.

It is now convenient to consider the variate $x=n\gamma y$ with density $h(x)$ given by

$$h(x)=c(x,n)\sum_{s=0}^{\infty}f_s(x)/\gamma^s, \quad f_0(x)=1,$$

where from (31) and the well-known result

$$\frac{1}{2\pi i}\int_{\epsilon-i\infty}^{\epsilon+i\infty}\frac{e^{\alpha y}\,d\alpha}{(1+\alpha/n\gamma)^{1/2N}}=\frac{n\gamma e^{-n\gamma y}(n\gamma y)^{1/2N-1}}{\left(\frac{1}{2}N-1\right)!}, \quad (y>0,\ N>0),$$

we have:

$$c(x,n)=\frac{e^{-x}x^{1/2(n-3)}}{\left(\frac{1}{2}(n-3)\right)!},$$

$$f_1(x)=\frac{B_2}{2}\left(n-\frac{1}{n}\right)\left(-1+\frac{x}{m}\right),$$

$$f_2(x)=\frac{B_2^2}{8}\left(n-\frac{1}{n}\right)^2\left(1-\frac{2x}{m}+\frac{x^2}{m(m+1)}\right),$$

$$f_3(x)=\frac{B_4}{4\cdot3}\left(n-\frac{1}{n^3}\right)\left(-1+\frac{x^3}{m(m+1)(m+2)}\right)$$
$$+\frac{B_2^3}{3!\,2^3}\left(n-\frac{1}{n}\right)^3\left(-1+\frac{3x}{m}-\frac{3x^2}{m(m+1)}+\frac{x^3}{m(m+1)(m+2)}\right), \quad (32)$$

and so on, where $m=\frac{1}{2}(n-1)$. We can now deduce that

$$E(\overset{*}{\gamma}/\gamma)^s=n^s\int_0^\infty c(x,n)\left(\sum\frac{f_s(x)}{\gamma^s}\right)\times\left\{\frac{1+\sqrt{(1+4x/3n\gamma)}}{4x}\right\}^s dx,\quad (s=1,2\ldots). \quad (33)$$

In particular, expanding the integrand to include the first two dominant terms, we have

$$E(\overset{*}{\gamma}/\gamma)^s\sim n^s\int_0^\infty\frac{e^{-x}x^{\frac{1}{2}(n-3)}}{\left(\frac{n-3}{2}\right)!(4x)^s}\left(1+\frac{f_1(x)}{\gamma}+\ldots\right)\times\left(2+\frac{2x}{3n\gamma}+\ldots\right)^s dx=\frac{n^s}{(n-3)(n-5)\ldots(n-2s-1)}\times\left\{1-\frac{s(s+1)}{3n\gamma}+\ldots\right\}. \quad (34)$$

This is exactly the same as the result for the maximum likelihood moment $E\,(\hat{\gamma}/\gamma)^s$ given in Shenton and Bowman (1969), so that the higher moments of Thom's statistic $\overset{*}{\gamma}$ are asymptotically equivalent to the corresponding moments of the maximum likelihood estimator of $\gamma$.

Similarly, by using the independence property given in section 6 in conjunction with (3b), an expression can be found for the asymptotic form of $E\,(\overset{*}{\beta}/\beta)^s$.

## REFERENCES

Abramowitz, Milton, and Stegun, Irene A., "Handbook of Mathematical Functions," *National Bureau of Standards Applied Mathematics Series 55*, Washington, D.C., **1964**, 1046 pp.

Cramer, Harold, *Mathematical Methods of Statistics*, Princeton University Press, N.J., **1946**, 575 pp.

de Bruijn, Nicholas G., *Asymptotic Methods in Analysis*, North-Holland Publishing Co., Amsterdam, **1961**, 200 pp.

Greenwood, J. Arthur, and Durand, David, "Aids for Fitting the Gamma Distribution by Maximum Likelihood," *Technometrics*, Vol. 2, No. 1, William Byrd Press, Richmond, Va., Feb. **1960**, pp. 55–65.

Masuyama, Motosaburo, and Kuroiwa, Y., "Table for the Likelihood Solutions of Gamma Distribution and Its Medical Applications," *Reports of Statistical Application Research*, Vol. 1, No. 1, Japan Publications Trading Co., Ltd., Tokyo, Mar. **1951**, pp. 18–23.

Shenton, L. R., and Bowman, K. O., "Properties of Estimators for the Gamma Distribution," Report CTC-1, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tenn., Sept. **1968**, 50 pp.

Shenton, L. R., and Bowman, K. O., "Maximum Likelihood Moments for the Two Parameter Gamma Distribution," *Sankhyā*, Vol. 31, Nos. 3 and 4, Statistical Publishing Society, Calcutta, **1969**, pp. 1–18.

Thom, Herbert C. S., "A Statistical Method of Evaluating Augmentation of Precipitation by Cloud Seeding," *Final Report, U.S. Advisory Committee on Weather Control*, Vol. 2, Washington, D.C., June **1957**, pp. 5–25.

Thom, Herbert C. S., "A Note on the Gamma Distribution," *Monthly Weather Review*, Vol. 86, No. 4, Apr. **1958**, pp. 117–122.

Thom, Herbert C. S., "Approximate Convolution of the Gamma and Mixed Gamma Distributions," *Monthly Weather Review*, Vol. 96, No. 12, Dec. **1968**, pp. 883–886.

[*Received September 5, 1969; revised October 17, 1969*]

# NEW ESSA PUBLICATIONS

## ESSA RESEARCH LABORATORIES

"An Introduction to Hydrodynamics and Water Waves:

"Volume I. Fundamentals," *ESSA Technical Report* ERL 118–POL 3–1, July 1969, 531 pp. Price: $3.75.

"Volume II. Water Wave Theories," *ESSA Technical Report* ERL 118–POL 3–2, July 1969, 240 pp. Price: $2.00.

These technical reports are based on lectures delivered to the Civil Engineering Department at Queen's University, Kingston, Ontario, in 1959–60. Their primary purpose is to present the foundations and essential aspects of the theoretical approach to hydrodynamics and water waves at a relatively simple level. A combination of both volumes can be considered as the text for a course in applied mathematics as well as the fundamentals of hydraulic and coastal engineering.

Volume I deals with the establishment of fundamental differential equations governing the flow motion in all possible cases. The possible approximations are also indicated. Methods of integrations and mathematical treatments of these equations are dealt with; integrations of general interest and integrations in some typical particular cases are presented.

Volume II is devoted to free surface flow motion and water wave theories.

---

These publications are available from Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.